Random Matrix Theory applied to the Estimation of Collision Multiplicities

Benoît Escrig
IRIT Laboratory
Université de Toulouse
Toulouse, France
E-mail: escrig@enseeiht.fr

***Abstract*—This paper presents two techniques in order to estimate the collision multiplicity, i.e., the number of users involved in a collision [1]. This estimation step is a key task in multi-packet reception approaches and in collision resolution techniques. The two techniques are proposed for IEEE 802.11 networks but they can be used in any OFDM-based system. The techniques are based on recent advances in random matrix theory and rely on eigenvalue statistics. Provided that the eigenvalues of the covariance matrix of the observations are above a given threshold, signal eigenvalues can be separated from noise eigenvalues since their respective probability density functions are converging toward two different laws: a Gaussian law for the signal eigenvalues and a Tracy-Widom law for the noise eigenvalues. The first technique has been designed for the white noise case, and the second technique has been designed for the colored noise case. The proposed techniques outperform current estimation techniques in terms of mean square error. Moreover, this paper reveals that, contrary to what is generally assumed in current multi-packet reception techniques, a single observation of the colliding signals is far from being sufficient to perform a reliable estimation of the collision multiplicities.**

***Index Terms*—multi-packet reception; collision multiplicity; model order selection; IEEE 802.11-based networks**

I. Introduction

Collisions are known to degrade the throughput of random access wireless networks, such as ad hoc networks[1]. A collision occurs when two or more user nodes access the channel simultaneously. Over the last decades, Medium Access Control (MAC) protocols have been designed with the rationale that all data packets from the colliding user nodes are lost when a collision occurs because the signals from all users mix.

It is is possible to retrieve part of the data packets that are involved in a collision, with approaches such as the capture effect [2]–[5]. Sophisticated techniques based on Multi-User Detection (MUD) allow the decoding of more than one data packets. MUD receivers have been successfully implemented in a wide range of application areas [6]. In this context, the number of colliding user is often needed to efficiently parameterize the MUD receivers.

Another approach is often invoked when a collision occurs. It consists in triggering a collision resolution (CR) mechanism where transmissions from the colliding users are re-scheduled in order to avoid another collision [2], [7]–[10]. In this context, CR mechanisms operate more efficiently when the number of colliding user is known. For instance, one could increase the contention window of a CR protocol with respect to the estimate of the number of colliding nodes. Note that our intention is not to propose a new CR mechanism but rather to provide a new parameter to improve the tuning of the mechanisms.

So, the purpose of this study is to estimate the number of colliding users, i.e., the collision multiplicity (CM) [1]. We focus here on the following scenario. The receiver at the destination node is processing a collision signal that consists of a mixture of signals from all users plus some Additive White Gaussian Noise (AWGN). The destination node is the node toward which all users involved in the collision are intending to send data [2]. So, from the observations of mixtures, the destination node has to estimate the number of original signals. This problem is a typical Model Order Selection (MOS) problem. MOS problems arise in the signal processing area and related areas such as signal array processing, radar, and sonar processing.

The MOS techniques are all based on the following rationale: the mixture of signals and noise can be decomposed into a noise subspace and a signal subspace, and the dimension of the signal subspace equals the number of signals. In order to perform this separation, the following steps are implemented. First, T observations of the mixture are gathered by the processing node. Then, the sample covariance matrix (SCM) of these observations is computed and an eigenvalue decomposition of this matrix is performed. These observations are obtained over T time-slots. The MOS techniques use the property that signal eigenvalues are much higher than noise eigenvalues, provided that the Signal to Noise Ratio (SNR) is high enough. When T eigenvalues are available and when K eigenvalues are significantly higher than the $T-K$ remaining eigenvalues, the conclusion is that there are K signals in the mixture.

The proposed approach has also been motivated by the following two observations: (i) many CM estimation techniques are based on the assumption that signal samples are

[1]In infrastructure-based networks, multiplexing techniques allow a fair bandwidth allocation among the users without any risk of collision.

[2]There are scenarios in which a collision occures between several source nodes transmitting toward several destination nodes. These scenarios are not addressed in this paper

uniformly distributed over a finite alphabet, i.e., signal samples are modulation symbols such as PSK or QAM symbols [11]–[13], and (ii) in many existing techniques, the number of observations is not much greater that the number of signals [7], [8].

Our point concerning (i) is to design estimation techniques that could be used in the context of Gaussian distributed samples. In this paper, we focus on the wireless standards that implement Orthogonal Frequency Division Multiplex (ODFM) transmissions; so signal samples can be considered as being Gaussian distributed and not uniformly distributed.[3] The issue that derives from (ii) is to clarify whether the number of observations can be made as small as $K+1$ or $K+2$ as in [7], [8] or not. Indeed, assuming that T and K are on the same order is in stark contrast with the typical assumptions that are used in signal processing [14]–[17].

We shall show that the proposed techniques outperform current techniques in terms of Mean Square Error (MSE) and that the number of observations T should be much higher than the number of signals K in order to obtain satisfactory MSE performance.

The rest of the paper is organized as follows. The introduction end ups with a Related Work subsection. The system model is introduced in Section II and some results on eigenvalue statistics are stated in Section III. The CMETs are described in Section IV. Simulation results are presented in Section V and a conclusion is drawn in the last section.

Related work

The first contributions, in the field of eigenvalue-based MOS techniques, were developed by Bartlett [18] and Lawley [19]. They propose a subjective setting of the threshold that is used to separate signal eigenvalues from noise eigenvalues. This approach is still used in some CR mechanisms [7], [8] in order to minimize the number of observations T. The algorithm starts with $T=1$ and the number of observations is incremented by one each time the smallest eigenvalue is higher than a noise threshold. As soon as the smallest eigenvalue crosses the noise threshold at step T, that means that the T^{th} highest eigenvalue is a noise eigenvalue, and hence that there are $T-1$ signals. Information theoretic criteria such as the Akaike Information Criterion (AIC) and the Minimum Description Length (MDL), developed by Wax and Kailath [17], have then been proposed in order to alleviate the limiting constraint imposed by the subjective setting of the separation threshold. The criteria are usually composed of a function that depends on the maximum likelihood estimator of the parameters of the model and a term that adjusts the first component to the context. This second term depends on the parameters of the system such as the number of samples per observation. The MDL have been used over several decades in several areas. An interesting review of this criterion can be found in [16].

[3]Note that signal samples in Code Division Multiple Access (CDMA) systems are also uniformly distributed since modulation symbols are multiplied by $+1/-1$ spreading codes.

The MDL is a consistent estimator of the number of signals when the number of observations, T, is fixed and when the number of samples per observations, denoted m, is such that $m \to +\infty$, provided that T is much larger than the number of signals K. Two limitations have recently been pointed out for this criterion.

First, the MDL have been shown to be inconsistent as the variance of the noise tends toward zero [15]. Second, the MDL is based on the distribution of the signal eigenvalues. It uses the property that sample eigenvalues, i.e., the eigenvalues that are obtained from the eigenvalue decomposition, are symmetrically distributed around the population eigenvalues, i.e., the theoretical eigenvalue. Basically, all the above-mentioned techniques are based on this property about signal eigenvalues. This centrality assumption makes sense in the T fixed, $m \to +\infty$ case. However, when T and m are on the same order, even when $T, m \to +\infty$, the previous property is no longer valid.

Recent advances in the Random Matrix Theory (RMT) field have brought into light several properties on the distribution of both signal and noise eigenvalues [20], [21]. These properties have been used in order to design new MOS techniques in the context where $T, m \to +\infty$ [16], [22]. These new techniques have been shown to outperform the classical MDL estimator.

Our purpose in this paper is to use these new properties in order to design new CM estimation techniques. Note that we are not in the context where $T, m \to +\infty$ but rather in the T fixed, $m \to +\infty$ case. The point is that current estimation techniques, whether they are based on the RMT or not, are based on the assumption that the number of observations T is much larger than the number of signals K. In our context, we want to minimize T with respect to K, so we are dealing with a context where T is in the order of K. Since the new RMT-based MOS techniques are performing in a wider range of parameter values, we believe that they can be considered as relevant candidates for our objective.

II. System Model

In this paper, we take in interest collision resolution algorithms and, more precisely, the estimation of the number of stations that are involved in a collision. We consider the scenario where K stations are simultaneously transmitting data packets to the same destination node (see Fig. 1).

We assume that the destination code and the colliding stations are all equipped with a single antenna. This is a worst case scenario. When the destination node is equipped with several antennas, observations are gathered more rapidly. The K colliding stations are transmitting OFDM frames of m samples each. The OFDM signal samples are Gaussian distributed with zero mean and variance unity. Moreover, the OFDM signals from the colliding users are assumed to be uncorrelated. This assumption makes sense since each transmission is affected by both a different Doppler shift and a different Doppler spread. The destination node receives T observations. So it is assumed that the destination can trigger transmissions from the colliding nodes. Note that T must be

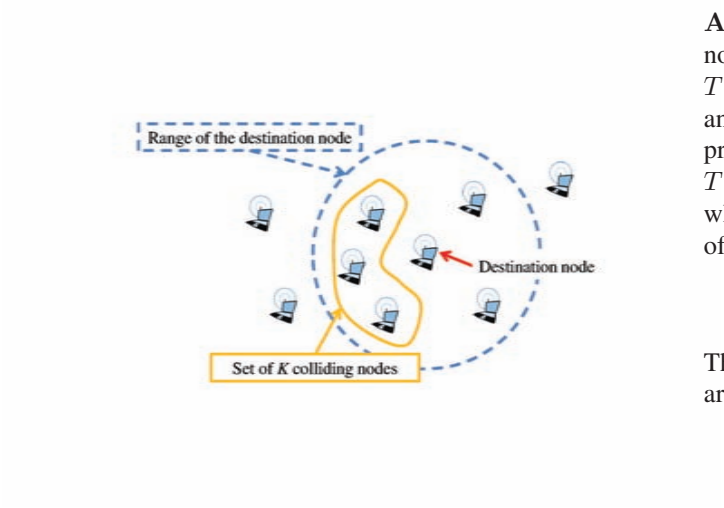

Fig. 1. Collision Scenario with $K = 3$ colliding nodes.

larger than K to allow an efficient estimation of K. This will be discussed in the simulation section. The destination triggers transmissions from the colliding users by sending a feedback frame. This frame serves also as a synchronization frame for user transmissions. Hence, the K users can be assumed to be coarsely synchronized in time. So this system model is similar to a source separation problem when K signals are impinging on a T sensor array. The $T \times 1$ observation vectors $\mathbf{y}_i$ can be written as

$$\mathbf{y}_i = \mathbf{A}\mathbf{x}_i + \mathbf{b}_i, \qquad i = 0, 1, \ldots, m \tag{1}$$

where $\mathbf{x}_i \sim \mathbb{C}\mathcal{N}_K(\mathbf{0}, \mathbf{R}_\mathbf{x})$ are $K \times 1$ complex Gaussian vectors with zero mean and covariance matrix $\mathbf{R}_\mathbf{x}$, $\mathbf{b}_i \sim \mathbb{C}\mathcal{N}_T(\mathbf{0}, \mathbf{R}_\mathbf{b})$ are $T \times 1$ complex noise vectors that are Gaussian distributed with zero mean and noise covariance matrix $\mathbf{R}_\mathbf{b}$. In the case of a white noise, we have that $\mathbf{R}_\mathbf{b} = \sigma^2\mathbf{I}_T$ where σ^2 denotes the noise variance and $\mathbf{I}_T$ is the $T \times T$ identity matrix. The channel matrix $\mathbf{A}$ is considered as an unknown $T \times K$ non-random matrix. For our study, we assume that the coefficients of $\mathbf{A}$ are modeled as circularly symmetric Gaussian coefficients with power unity (Rayleigh fading). The channels gains are assumed to have constant values over the duration of the frame and change randomly from one frame to another. This corresponds to the typical block-fading channel assumption.

The observations $\mathbf{y}_i$ can be whitened by the following transformation

$$\mathbf{y}_i^\dagger = \mathbf{R}_\mathbf{b}^{-1/2}\mathbf{y}_i$$

provided that the noise covariance matrix $\mathbf{R}_\mathbf{b}$ is known *a priori* and is nonsingular, where $\mathbf{R}_\mathbf{b}^{+1/2}$ is the Hermitian nonnegative definite square root of $\mathbf{R}_\mathbf{b}$. The transformation simply reduces to a normalization step in the case of a white Gaussian noise. The covariance matrix $\mathbf{R}$ of the snapshots $\mathbf{y}_i$ is given by

$$\mathbf{R} = \mathbf{A}\mathbf{R}_\mathbf{x}\mathbf{A}^\mathrm{H} + \mathbf{R}_\mathbf{b} = \Psi + \mathbf{R}_\mathbf{b}$$

with $^\mathrm{H}$ denoting the complex conjugate, the signal and noise vectors being independent. We assume that the channel matrix $\mathbf{A}$ is full rank and that the signal covariance matrix $\mathbf{R}_\mathbf{x}$ is nonsingular so that the rank of Ψ is $\min(K, T)$. Hence, when $T \leq K$, there are T non-zero eigenvalues in the matrix Ψ and when $T > K$, there are K non-zero eigenvalues. This property is used in [7], [8] where the number of observations, T, is incremented until $\mathbf{R}$ has zero eigenvalues. When the whitening transformation is applied, the covariance matrix $\mathbf{R}^\dagger$ of the whitened observations, denoted $\mathbf{R}^\dagger$, is defined as

$$\mathbf{R}^\dagger = \mathbf{R}_\mathbf{b}^{-1/2}\mathbf{R}\mathbf{R}_\mathbf{b}^{-1/2} = \mathbf{R}_\mathbf{b}^{-1/2}\Psi\mathbf{R}_\mathbf{b}^{-1/2} + \mathbf{I}_T$$

The population eigenvalues $\mathbf{R}^\dagger$, denoted $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_T$, are such that

$$\lambda_i > 1 \quad \text{for } 1 \leq i \leq \min(K, T) \tag{2}$$

$$\lambda_i = 1 \quad \text{for } \min(K, T) < i \leq T \tag{3}$$

In most approaches, T is much larger that K so the conditions in (2) and (3) should be written as $1 \leq i \leq K$ and $K < i \leq T$ respectively. However, in some CR algorithms such as the ones in [7] and in [8], (2) and (3) are used in this way, since they are initialized with $T = 1$ and T is incremented by one each time the T^{th} smallest eigenvalue is not detected has being a noise eigenvalue. In any case, the estimation of K can be performed from the multiplicity of the λ_i equalling one. When $\mathbf{R}$ and $\mathbf{R}_\mathbf{b}$ are known, and when the rank of $\mathbf{R}_\mathbf{b}^{-1}\Psi$ is K, the CM estimation can be easily performed from the multiplicity of the λ_i equalling one. Otherwise, when $\mathbf{R}$ and $\mathbf{R}_\mathbf{b}$ are unknown and have to be estimated, the sample covariance matrix (SCM) of the observations $\mathbf{y}_i$, denoted $\widehat{\mathbf{R}}$, must be computed

$$\widehat{\mathbf{R}} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{y}_i\mathbf{y}_i^\mathrm{H}$$

The SCM of the noise, denoted $\widehat{\mathbf{R}}_\mathbf{b}$, must be also computed using

$$\widehat{\mathbf{R}}_\mathbf{b} = \frac{1}{N}\sum_{j=1}^{N}\mathbf{b}_j\mathbf{b}_j^\mathrm{H}$$

Note that the $\mathbf{b}_j$, $1 \leq j \leq N$ are independent noise-only samples. We assume that we can get noise-only samples by using idle time slots. In that case, the estimation of K is based on the eigenvalue decomposition of the SCM $\widehat{\mathbf{R}}^\dagger$

$$\widehat{\mathbf{R}}^\dagger = \frac{1}{m}\sum_{i=1}^{m}\mathbf{y}_i^\dagger(\mathbf{y}_i^\dagger)^\mathrm{H} = \widehat{\mathbf{R}}_\mathbf{b}^{-1/2}\widehat{\mathbf{R}}\widehat{\mathbf{R}}_\mathbf{b}^{-1/2} \tag{4}$$

The sample eigenvalues of the SCM matrix $\widehat{\mathbf{R}}^\dagger$ are denoted $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \ldots \geq \hat{\lambda}_T$.

III. Some Results on Eigenvalue Statistics

In the following, we investigate new results in RMT. In particular, we provide new characterizations of the eigenvalues of SCMs. We first begin with the signal-free case and then address the signal bearing case.

A. Signal-free Case

We start with the signal-free case where no user is transmitting, so $K = 0$. As a consequence, we have that

$$\begin{aligned} \mathbf{R} &= \mathbf{R_b} \\ \mathbf{R}^{\dagger} &= \mathbf{I}_T \end{aligned} \tag{5}$$

and all population eigenvalues λ_i are all equal to one. From the system model and (5), the vectors $\mathbf{y}_i$ are T-dimensional complex Gaussian independent vectors such that $\mathbf{y}_i \sim \mathbb{C}\mathcal{N}_T(\mathbf{0}, \mathbf{R_b})$.

From Property 6.1, we get that $m\widehat{\mathbf{R}}$ is a T-variate complex Wishart matrix with m degrees of freedom and covariance matrix $\mathbf{R_b}$, i.e.,

$$m\widehat{\mathbf{R}} \sim \mathbb{C}W_T(m, \mathbf{R_b})$$

Similarly, since we have that $\mathbf{y}_i^{\dagger} \sim \mathbb{C}\mathcal{N}_T(\mathbf{0}, \mathbf{I}_T)$, we have that

$$m\widehat{\mathbf{R}}^{\dagger} \sim \mathbb{C}W_T(m, \mathbf{I}_T)$$

From Property 6.2, we now characterize the eigenvalues of $\widehat{\mathbf{R}}$ in the signal-free case and in the presence of a white Gaussian noise [16], [20], [23], [24].

Corollary 3.1: In the signal-free case, when $\mathbf{y}_i \sim \mathbb{C}\mathcal{N}_T(\mathbf{0}, \lambda\mathbf{I}_T)$, the Empirical Spectral Distribution (ESD) of the $T \times T$ random matrix $\widehat{\mathbf{R}}^{\dagger}$ converges almost surely to the Marčenko-Pastur law in (14) (see Fig. 2).

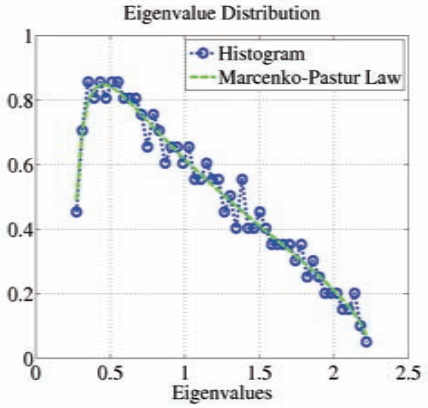

Fig. 2. Distribution of the eigenvalues of a matrix $\widehat{\mathbf{R}}$ in the signal-free case with $T = 500$ and $m = 2000$ ($\lambda = 1$).

When the snapshots $\mathbf{y}_i$ are whitened and become $\mathbf{y}_i^{\dagger}$, a new distribution for the ESD must be considered [22].

From Property 6.3, we now characterize the eigenvalues of $\widehat{\mathbf{R}}^{\dagger}$ in the signal-free case in the presence of an arbitrarily colored Gaussian noise [20], [22]–[24].

Corollary 3.2: In the signal-free case, when $\mathbf{y}_i^{\dagger} \sim \mathbb{C}\mathcal{N}_T(\mathbf{0}, \mathbf{I}_T)$, the ESD of the $T \times T$ random matrix $\widehat{\mathbf{R}}^{\dagger}$ converges almost surely to the modified Marčenko-Pastur law in (16) (see Fig. 3).

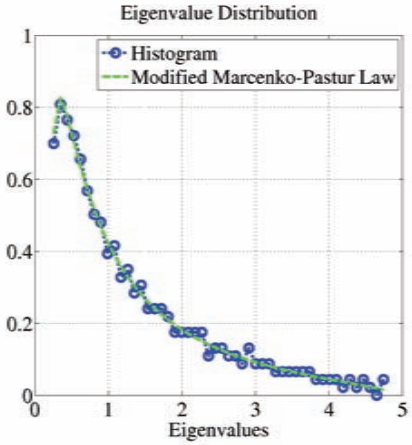

Fig. 3. Distribution of the eigenvalues of a matrix $\widehat{\mathbf{R}}^{\dagger}$ in the signal-free case with $T = 500$ and $m = 2000$.

In the signal-free case, we are interested in the eigenvalues of $\widehat{\mathbf{R}}^{\dagger}$ or, equivalently, in the eigenvalues θ that satisfy

$$\widehat{\mathbf{R}}v = \theta\widehat{\mathbf{R}}_{\mathbf{b}}v \tag{6}$$

We can rewrite (6) as follows

$$m\widehat{\mathbf{R}}v = (\frac{m}{N}\theta)N\widehat{\mathbf{R}}_{\mathbf{b}}v \tag{7}$$

Since $m\widehat{\mathbf{R}} \sim \mathbb{C}W_T(m, \mathbf{R_b})$ and $N\widehat{\mathbf{R}}_{\mathbf{b}} \sim \mathbb{C}W_T(N, \mathbf{R_b})$, (7) is similar to (20). So, using Property 6.5, the largest eigenvalue $\hat{\lambda}_1$ that satisfies (7) is Tracay-Widom (TW) distributed and we can derive the following property.

Property 3.1: The largest eigenvalue $\hat{\lambda}_1$ that satisfies (6) is TW distributed, i.e.,

$$\mathrm{P}\left\{\frac{\log(\frac{m}{N}\hat{\lambda}_1) - \mu(T,m,N)}{\sigma(T,m,N)} \leq x\right\} \rightarrow TW_{\mathbb{C}}(x)$$

The pdf of the largest eigenvalue $\hat{\lambda}_1$ is represented in Fig. 4. Note that this characterization uses explicitly the double Wishart setting that has been motivated by the need to whiten the observations when the additive Gaussian noise of the channel is not white. Another and simpler characterization of $\hat{\lambda}_1$ has also be proposed in [20], [21].

Property 3.2: In the signal-free case, the whitened snapshots $\mathbf{y}_i^{\dagger}$ are $\mathcal{N}_T(\mathbf{0}, \mathbf{I}_T)$ and the largest eigenvalue $\hat{\lambda}_1$ of the SCM $\widehat{\mathbf{R}}^{\dagger}$ is Tracy-Widom distributed. When $T, m \rightarrow \infty$ such that $T/m \rightarrow c \in (0, \infty)$,

$$\mathrm{P}\left[\frac{m\hat{\lambda}_1 - \mu_{T,m}}{\sigma_{T,m}} \leq x\right] \rightarrow TW_{\mathbb{C}}(x)$$

where

$$\begin{aligned} \mu_{T,m} &= (\sqrt{T} + \sqrt{m})^2 \\ \sigma_{T,m} &= (\sqrt{T} + \sqrt{m})(\frac{1}{\sqrt{T}} + \frac{1}{\sqrt{m}})^{1/3} \end{aligned}$$

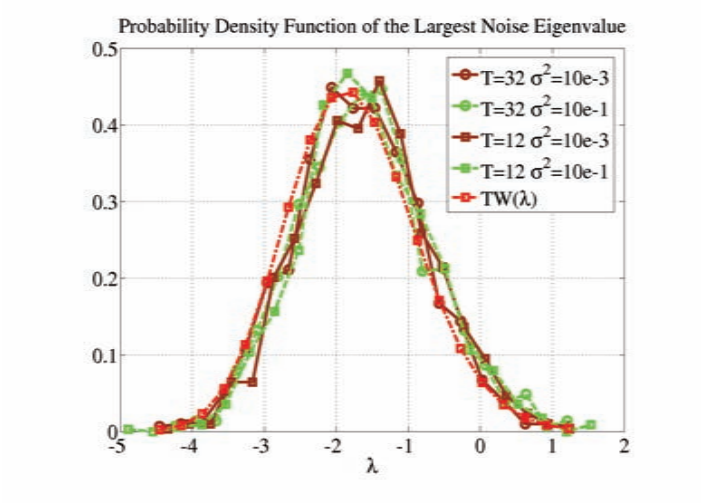

Fig. 4. Probability density function of the largest eigenvalue of a matrix $\widehat{\mathbf{R}}^{\dagger}$ in the signal-free case with $m = 61440$ and different values for T and the noise variance σ^2.

When the snapshots $\mathbf{y}_i^{\dagger}$ are $\mathcal{N}_T(\mathbf{0}, \sigma^2\mathbf{I}_T)$, the convergence limit of $m\hat{\lambda}_1$ is $\sigma^2(\sqrt{T}+\sqrt{m})^2$. This corresponds to the non-normalized case. Note that the convergence rate to the $TW_{\mathbb{C}}(x)$ distribution function is $\mathcal{O}(T^{-1/3})$. When the parameters m and T are not so large, which is practically the case when we want to reduce the number of observations, the convergence rate to the TW distribution is rather $\mathcal{O}(T^{-2/3})$ provided that the mean and standard deviation have been modified appropriately .

B. Signal Bearing Case

When there are K signals and when $T \to \infty$, the limiting ESD of $\widehat{\mathbf{R}}^{\dagger}$ still converges to a Marčenko-Pastur distribution. This is because all eigenvalues are equally weighted by the ESD so the impact of K signals vanishes when $T \to +\infty$. Now there are K signal eigenvalues and $T-K$ noise eigenvalues, and their respective characterization are different.[4] Noise eigenvalues are still distributed according to a TW distribution. The situation is a little bit more complicated for the signal eigenvalues since the characterization depends on a threshold τ. In the case of a white Gaussian noise, when no whitening transfomation in applied, the threshold is defined as $\tau = 1+\sqrt{c}$ [16]. The threshold takes into account the c_1 ratio in Property 6.3, when dealing with a colored noise [22]. When the i^{th} largest signal eigenvalue $\hat{\lambda}_i$ is strictly higher than τ, the signal eigenvalue converges to a limit different from that in the signal-free case [16]. In that case, the signal eigenvalue is Gaussian distributed, i.e.,

$$\mathrm{P}\left[\frac{\hat{\lambda}_i - \mu_i}{\sigma_i} \leq x\right] \to G(x)$$

[4]This case is often referred to as a "spiked" model in RMT.

where

$$\begin{aligned}\mu_i &= \lambda_i(1+\frac{c}{\lambda_i-1})\\ \sigma_i &= \frac{\lambda_i}{\sqrt{T}}\sqrt{1-\frac{c}{(\lambda_i-1)^2}}\end{aligned}$$

and

$$G(x) = \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y^2}{2}\right)dy$$

In Fig. 5, the pdf of the smallest signal eigenvalue is represented. Otherwise, when the i^{th} largest signal eigenvalue $\hat{\lambda}_i$

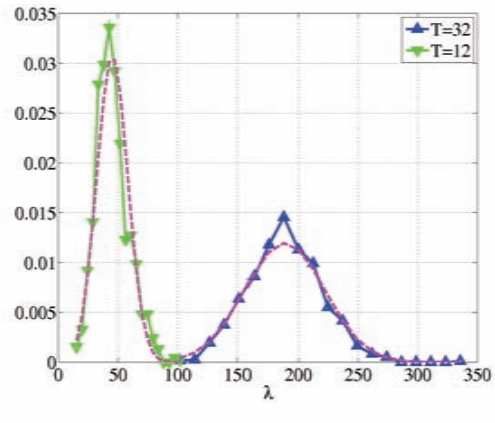

Fig. 5. Probability density function of the smallest signal eigenvalue of a matrix $\widehat{\mathbf{R}}^{\dagger}$ in the signal-bearing case with $K = 4$, i.e. $\hat{\lambda}_5$, with $m = 61440$, a noise variance of $\sigma^2 = 0.1$, and different values for T. The solid curves with markers are the simulation results and the dotted curves are the Gaussian distribution.

is strictly lower that τ, the signal eigenvalue is distributed according to a TW distribution, i.e.,

$$\mathrm{P}\left\{\frac{\log(\frac{m}{N}\hat{\lambda}_i) - \mu(T-i,m,N)}{\sigma(T-i,m,N)} \leq x\right\} \to TW_{\mathbb{C}}(x)$$

Note that there is a specific phenomenon when the i^{th} largest signal eigenvalue exactly equals to the threshold [24]. Hence, when $K \ll T$, the signal eigenvalues exhibit, on rescaling, fluctuations described by the TW distributions when there are strictly below the threshold τ. So the distributions obtained for the signal-free case ($K = 0$) closely approximate the distribution of the signal eigenvalues. These results have an impact on the design of CM estimation techniques since signal eigenvalues strictly below τ are considered as being noise eigenvalues. Note that adding observations, i.e., increasing T, cannot be an option to tackle this problem since τ grows with T. However, when the signal eigenvalues are above the threshold, a reliable estimation of the CM is possible.

IV. Collision Multiplicity Estimation Techniques

We first present the CM estimation method based on the distribution of noise eigenvalues. Then we review two approaches based on the distribution of signal eigenvalues.

A. A method based on the distribution of noise eigenvalues

We assume that T observations are available at the destination node. The SCM $\widehat{\mathbf{R}}^{\dagger}$ is computed from (4). Then, an eigenvalue decomposition of $\widehat{\mathbf{R}}^{\dagger}$ is performed and the resulting eigenvalues are sorted in descending order $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq ... \geq \hat{\lambda}_T$. The method relies on the property that the i^{th} largest noise eigenvalue is TW distributed, provided that all signal eigenvalues are located above the detectability threshold. So for every $l = 0, 1, \ldots, T-1$, we test the null hypothesis that "there are exactly l colliding signals" against the alternative hypothesis that there are "at least $l+1$ colliding signals". The test is performed by computing the test statistic $S_\lambda(l)$

$$S_\lambda(l) = \frac{\log(\frac{m}{N}\hat{\lambda}_{l+1}) - \mu(T-l, m, N)}{\sigma(T-l, m, N)} \tag{8}$$

and comparing it to a threshold τ_α . The threshold τ_α is defined as

$$\tau_\alpha = TW_{\mathbb{C}}^{-1}(1-\alpha) \tag{9}$$

where α is some significance level. More precisely, a significance level α is set so that the probability that the null hypothesis is detected by chance is α. So the null hypothesis is accepted when we have

$$S_\lambda(l) < \tau_\alpha \tag{10}$$

The sequence of tests begins with $l = 0$. If the test $S_\lambda(0)$ is passed, i.e., when $S_\lambda(0) < \tau_\alpha$, there are no colliding users. Otherwise, there is at least one signal. The procedure proceeds to the next step with $l = 1$. The tests are performed subsequently until the test is passed. When the test is passed at step $l = q$, the eigenvalue $\hat{\lambda}_{q+1}$ is detected as being TW distributed. So $\hat{\lambda}_{q+1}$ is a noise eigenvalue, and hence there are exactly q colliding signals. Once the number of colliding users has been determined, the destination node stops the CM estimation process and proceeds to the next processing block. The procedure is summarized in Algorithm 1. The

Algorithm 1 TWIT algorithm

```
Compute R̂†
Perform the eigenvalue decomposition of R̂†
Sort the eigenvalues λ̂_i, i = 1,...,T of R̂†
K̂_TWIT ← 0 and Test ← False
while Test = False and K̂_TWIT < T do
   μ ← μ(T − K̂_TWIT, m, N)
   σ ← σ(T − K̂_TWIT, m, N)
   Test ← {σ^{-1}[log(mλ̂_{K̂_TWIT+1}/N) − μ] < τ_α}
   if Test = False then
      K̂_TWIT ← K̂_TWIT + 1
   else
      break
   end if
end while
```

mean $\mu(x, y, z)$ and the standard deviation $\sigma(x, y, z)$ in the algorithm are defined in Property 6.4. Note that this criterion has been originally designed for arbitrary (or colored) noise [22]. That is the reason why the algorithm uses the eigenvalues of $\widehat{\mathbf{R}}^{\dagger}$.

B. Methods based on the distribution of signal eigenvalues

We present two CMETs that are all based on the Gaussian distribution of signal eigenvalues. We first present the well-known MDL criterion, and then address another criterion based on recent advances in RMT.

1) The MDL criterion: The MDL criterion has been defined in [17]. This criterion has been used for decades in the area of signal array processing, and other related domains. The MDL estimator $\hat{K}_{\text{MDL}}$ for the CM K is defined as

$$\hat{K}_{\text{MDL}} = \underset{k=1,\ldots,T}{\operatorname{argmin}} \{\text{MDL}(k)\}$$

where

$$\text{MDL}(k) = -m(T-k)\log\left[\frac{g(k)}{a(k)}\right] + \frac{1}{2}k(2T-k)\log(m)$$

where

$$g(k) = \prod_{i=k+1}^{T} \hat{\lambda}_i^{\frac{1}{T-k}} \qquad a(k) = \frac{1}{T-k}\sum_{i=k+1}^{T}\hat{\lambda}_i$$

where the $\hat{\lambda}_i$ denote the eigenvalues of $\widehat{\mathbf{R}}^{\dagger}$ with $1 \leq i \leq T$, ordered in descending order. This estimator is consistent in the $m \to \infty$ sense. One of the reason why the MDL criterion has been widely used over the past two decades comes from its robustness to model mismatch, in particular when the underlying assumptions of snapshots and noise Gaussianity can be relaxed [25]–[27].

2) The criterion based on recent results on signal eigenvalues: This new criterion has been designed using the last results in RMT [16]. We denote this estimator SEMOS (Signal Eigenvalue-based Model Order Selection). The estimator $\hat{K}_{\text{SEMOS}}$ for the CM K is defined as, for complex data

$$\hat{K}_{\text{SEMOS}} = \underset{k=0,\ldots,T-1}{\operatorname{argmin}} \left\{ \left[\frac{1}{2}(\frac{1}{c})^2 q_k^2\right] + 2(k+1) \right\}$$

The test statistic q_k is defined as

$$q_k = \left[(T-k)\frac{g_s(k)}{a_s(k)} - (1+c)\right] \times T \tag{11}$$

where

$$g_s(k) = \sum_{i=k+1}^{T}\hat{\lambda}_i^2 \qquad a_s(k) = (\sum_{i=k+1}^{T}\hat{\lambda}_i)^2$$

V. Simulation Results

We study the performance of three estimation techniques: the MDL criterion, the TWIT, and the SEMOS estimator. The methods are evaluated in the context of Rayleigh block-fading channels. The channel coefficients A_{ij} in (1) are circularly symmetric Gaussian coefficients with zero mean and power unity and the coefficients are randomly changing from one observations to another. User stations are transmitting OFDM signals. The signals are constructed according to the

TABLE I
OFDM SIGNAL PARAMETERS

Modulation	BPSK
Number of sub-carriers N_{sub}	1024
Guard Interval GI	1/4
Number of OFDM symbols per OFDM block N_{OFDM}	48

IEEE 802.11 standard [28] and use the signal parameters listed in Table I.

From the signal parameters, we get that the number of samples per OFDM frame is $m = (1 + \text{GI}) \times N_{\text{sub}} \times N_{\text{OFDM}} = 61440$. For the sake of simplicity, we set that $N = m$, i.e., $N = 61440$. The performance of estimation techniques have been evaluated over 10,000 Monte Carlo trials ($N_{\text{Simu}} = 10000$) using the MATLAB software.

The three estimation techniques are compared in Fig. 6 to 10 in the case of a white Gaussian noise. The results have been obtained for both a variable T and a variable SNR. The significance level α for the TWIT is set to 0.01 [22]. Figures 6 and 7 represent the simulation results for $K = 3$ and $K = 4$ respectively. The simulation results are similar with $K = 3$ and $K = 4$ so we focus on the latter case in the following. When the number of observations and the SNR are increasing, the curves are converging toward the same position so it is difficult to compare them. So we choose instead to compare the estimation techniques with respect to the MSE ϵ^2 between the estimates $\hat{K}$ and K

$$\epsilon^2 = \frac{1}{N_{\text{Simu}}} \sum_{i=1}^{N_{\text{Simu}}} |\hat{K}_i - K|^2 \tag{12}$$

where $\hat{K}_i$ is the estimate of K for the i^{th} trial. This is shown in Fig. 8 to 10. Since the curves are plotted on a log scale, only non-zero values are plotted. So when the experimental MSE is zero, the corresponding point of the curve is not represented.

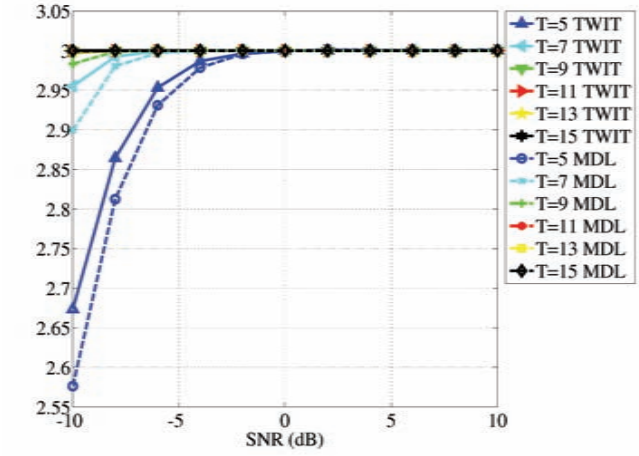

Fig. 6. Estimates of $K = 3$ with the TWIT and the MDL criterion for a variable number of snapshots T, $5 \leq T \leq 15$, and different SNR values.

The simulation results show that the estimation techniques perform better at high SNR and when T increases. The number of observations should be higher than almost three times the CM K in order to get a relative MSE lower than 10%. Similarly, the SNR should be higher than 0 dB in order to get a similar performance level, for any values of T. The two RMT-based techniques, i.e., the SEMOS estimator and the TWIT, outperforms the MDL criterion. However, the differences between the simulation curves decrease when the SNR or T is increasing. Moreover, the SEMOS estimator outperforms the TWIT. This can be due to the fact that the former technique used all the eigenvalues in the computation of the estimator while the latter only uses one eigenvalue to detect the number of signals. Note, however, that the TWIT has not been designed with the purpose of outperforming the SEMOS technique. Rather, the TWIT has been designed to cope with colored Gaussian noises. This is illustrated in the next figure.

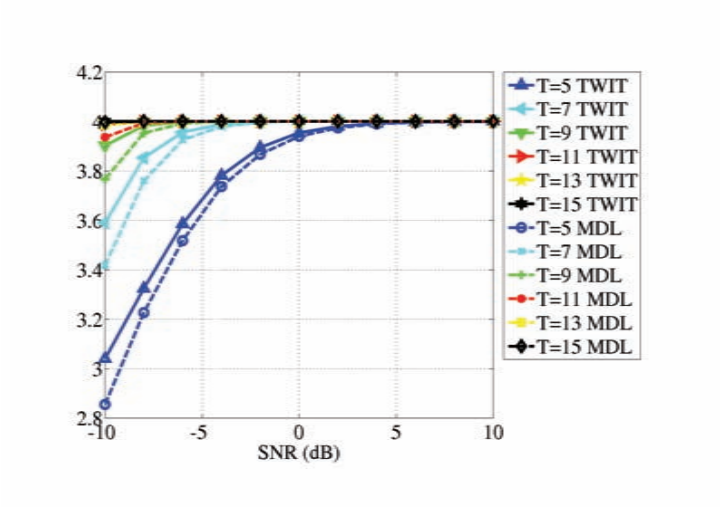

Fig. 7. Estimates of $K = 4$ with the TWIT and the MDL criterion for a variable number of snapshots T, $5 \leq T \leq 15$, and different SNR values.

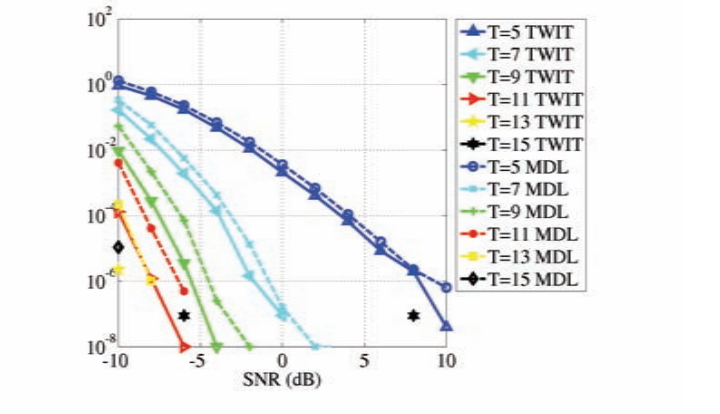

Fig. 8. MSE of two estimation techniques (TWIT and the MDL criterion) for $K = 4$, a variable number of snapshots T, $5 \leq T \leq 15$, and different SNR values.

Fig. 11 represents the performance of both the SEMOS

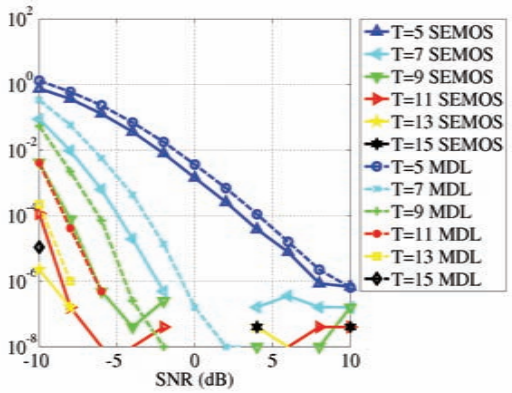

Fig. 9. MSE of two estimation techniques (the SEMOS estimator and the MDL criterion) for $K = 4$, a variable number of snapshots T, $5 \leq T \leq 15$, and different SNR values.

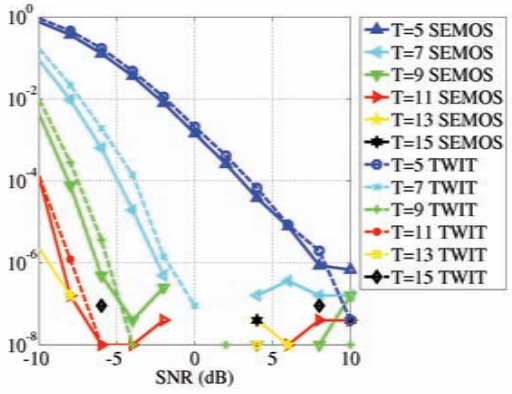

Fig. 10. MSE of two estimation techniques (the SEMOS estimator and the TWIT) for $K = 4$, a variable number of snapshots T, $5 \leq T \leq 15$, and different SNR values.

estimator and the TWIT. The simulation parameters are the same as the ones that have been used in previous simulations. The noise is now a colored noise. For this purpose, a white Gaussian noise with variance unity is passed through a filter with N_f coefficients f_i, $1 \leq i \leq N_f$ given in Table II.

We have that

$$\sum_{i}^{N_f} |f_i|^2 = 1$$

so the colored Gaussian noise at the output of the filter is also unit variance. The simulation results in Fig. 11 show that the SEMOS estimator is unable to estimate correctly the CM while the TWIT is still performing well.

TABLE II
FILTER COEFFICIENTS

0.227	0.460	0.688	0.460	0.227

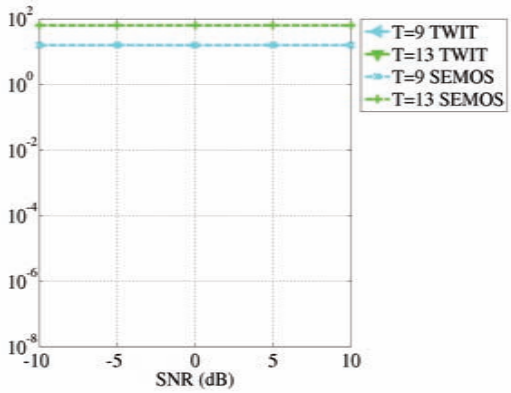

Fig. 11. MSE of two estimation techniques (the SEMOS estimator and the TWIT criterion) for $K = 4$, a variable number of snapshots T, $5 \leq T \leq 15$, and different SNR values, in the presence of a colored noise.

A. Discussion

The simulation results show that the estimation techniques perform better when the number of observations T is much larger than the number of signals K. This is in stark contrast with previous studies, such as the ones implemented in [7] and [8] where T is set to a value not greater than $K + 2$. In other approaches [11], [12], a single observation is required to perform the MUD of the colliding users. These references do not explicitly claim that the number of signals is estimated using only one observation. In this context, the blind separation technique designed in [13], is implemented and the number of sources (users) is assumed to have been estimated using an MDL-like criterion.

VI. CONCLUSION

In this paper, two new CM estimation techniques, the TWIT and the SEMOS estimator, have been proposed. The methods are based on eigenvalue statistics. In the TWIT, eigenvalues are tested in descending order, from the largest to the lowest. The first eigenvalue $\hat{\lambda}_q$ that is considered as being Tracy-Widom distributed allows the CM estimation by $\hat{K} = q - 1$. In the SEMOS estimator, a criterion is computed just as in the MDL criterion approach. These CM estimation techniques have been shown to outperform the typical MDL criterion. The SEMOS estimator outperforms the TWIT in the presence of a white Gaussian noise but it is inefficient in the presence of a colored noise. Moreover, simulation results have shown that a large number of snapshots T is needed in order to allow a good estimation of K in terms of MSE. Furthermore, the number of snapshots must be significantly higher than the number of colliding users K ($T \gg K$). These settings are similar to the settings that are used in MOS techniques for signal array processing.

The impact of these results is twofold. First, some CR techniques such as the network-assisted diversity multiple access (NDMA) [7], [8] cannot be implemented in IEEE 802.11

networks notably because these CR techniques are based on the assumption that T can be made as small as $K+1$ or $K+2$. Second, some multiple packet reception (MPR) protocols for IEEE 802.11 networks that use the blind user separation in [13] appear to be rather questionable since they assume that the CM estimation can rely on a single observation of collided request-to-send (RTS) frames. Even if the destination node is equipped with four antennas ($T = 4$), our simulation results have shown that the receiver at the destination node needs many more snapshots in order to provide a good estimation of K. This paper has pointed out a strong constraint in the design of MPR techniques. It revealed that a single observation of the colliding signals is far from providing enough information to estimated the number of colliding nodes.

Further investigations are now needed in order to fully characterized the performance of the proposed CM estimation techniques in typical operating conditions. The obtained results will allow the implementation of these techniques in current or future standards.

APPENDIX

A. Complex Wishart Matrices

Property 6.1: Consider a $T \times m$ matrix $\mathbf{X}$ with m samples drawn from a T-dimensional complex Gaussian law with zero mean and covariance matrix $\mathbf{R}_{\mathbf{X}}$, denoted $\mathbb{C}\mathcal{N}_T(\mathbf{0}, \mathbf{R}_{\mathbf{X}})$. The random matrix $\mathbf{X}\mathbf{X}^{\mathrm{H}}$ is a T-variate complex Wishart matrix with m degrees of freedom [20]. This is denoted

$$\mathbf{X}\mathbf{X}^{\mathrm{H}} \sim \mathbb{C}W_T(m, \mathbf{R}_{\mathbf{X}})$$

where $\mathbb{C}W_T(m, \mathbf{R}_{\mathbf{X}})$ denotes complex Wishart law, parameterized accordingly.

B. Empirical Spectral Distribution (ESD) of a Matrix

The ESD of an $T \times T$ Hermitian matrix $\mathbf{X}$, denoted $F^{\mathbf{X}}(x)$, is the distribution function of the eigenvalues of $\mathbf{X}$, i.e., for $x \in \mathbb{R}$

$$F^{\mathbf{X}}(x) = \frac{1}{T} \sum_{i=1}^{T} 1_{\{x, \lambda_i \leq x\}}(x) \tag{13}$$

where $\lambda_1, \ldots, \lambda_T$ are the population eigenvalues of $\mathbf{X}$ and $1_A(x)$ is the indicator function of the set A, i.e., $1_A(x) = 1$ if $x \in A$, and $1_A(x) = 0$ otherwise. The Hermitian property is required to ensure that all eigenvalues of $\mathbf{X}$ belong to the real line.

Property 6.2: Let $\mathbf{X} \in \mathbb{C}^{T \times m}$ be a random matrix with independent and identically distributed (iid) entries X_{ij} such that X_{ij} has zero mean and variance λ. As $T, m \to \infty$ with $T/m \to c \in (0, \infty)$, the ESD of

$$\mathbf{B} = \frac{1}{m} \mathbf{X}\mathbf{X}^{\mathrm{H}}$$

converges almost surely to a non-random distribution function with density $f_c(x)$ given by

$$\begin{aligned} f_c(x) &= (1 - \frac{1}{c})^+ \times \delta(x) \\ &+ \frac{1}{2\pi\lambda x c} \sqrt{(x-a)^+ (b-x)^+} \end{aligned} \tag{14}$$

with $a = (1 - \sqrt{c})^2$, $b = (1 + \sqrt{c})^2$, $\delta(x) = 1_{\{0\}}(x)$ and, for $x \in \mathbb{R}$, $x^+ = \max(0, x)$.
The probability density function (pdf) $f_c(x)$ is the Marčenko-Pastur density.

Property 6.3: Let $\mathbf{B} \in \mathbb{C}^{T \times m}$ be a random matrix defined as

$$\mathbf{B} = \frac{1}{m} \mathbf{T}^{1/2} \mathbf{X}\mathbf{X}^{\mathrm{H}} \mathbf{T}^{1/2} \tag{15}$$

where $\mathbf{X} \in \mathbb{C}^{T \times m}$ is a random matrix with iid entries that have zero mean and power unity, and where $\mathbf{T}$ is an $T \times T$ Hermitian non negative definite matrix. As $T, m \to \infty$ with $T/m \to c \in (0, \infty)$, the ESD of $\mathbf{B}$ converges almost surely to a non-random distribution function with density $\widetilde{f}_c(x)$ given by

$$\begin{aligned} \widetilde{f}_c(x) &= (1 - \frac{1}{c})^+ \times \delta(x) \\ &+ \frac{1 - c_1}{2\pi x (x c_1 + c)} \sqrt{(x - b_1)^+ (b_2 - x)^+} \end{aligned} \tag{16}$$

with

$$\begin{aligned} b_1 &= \left[\frac{1 - \sqrt{1 - (1-c)(1-c_1)}}{1 - c_1} \right]^2 \\ b_2 &= \left[\frac{1 + \sqrt{1 - (1-c)(1-c_1)}}{1 - c_1} \right]^2 \end{aligned}$$

where $T/N \to c_1 \in (0, 1)$ as $T \to +\infty$.
The probability density function (pdf) $\widetilde{f}_c(x)$ is a modified Marčenko-Pastur density.

C. Distribution of the Largest Eigenvalue of Wishart Matrices

We now investigate the distribution of the largest eigenvalue of Wishart matrices. Consider $\mathbf{A} \sim \mathbb{C}W_T(m, \mathbf{R}_{\mathbf{b}})$ and $\mathbf{B} \sim \mathbb{C}W_T(N, \mathbf{R}_{\mathbf{b}})$, two independent random matrices where $m, N > T$. We consider the generalized eigenproblem constructed from $\mathbf{A}$ and $\mathbf{B}$

$$\mathbf{A}v = \theta(\mathbf{A} + \mathbf{B})v \tag{17}$$

where v denotes the eigenvector corresponding to the eigenvalue θ. This is referred to as a double Wishart setting [20].

Property 6.4: The largest eigenvalue satisfying (17), denoted $\lambda_1^{(D)}$, is distributed according to a Tracy-Widom (TW) distribution when $m, N \to \infty$ as $T \to \infty$ with $m, N > T$, i.e.,

$$\mathrm{P}\left[\frac{W(\lambda_1^{(D)}) - \mu(T, N, m)}{\sigma(T, N, m)} \leq x \right] \to TW_{\mathbb{C}}(x) \tag{18}$$

where $W(\theta)$ denotes the logit transformation of θ, i.e., $W(\theta) = \log[\theta/(1-\theta)]$ and $TW_{\mathbb{C}}(x)$ is the TW distribution function for complex data.

The centering $\mu(T,N,m)$ and rescaling $\sigma(T,N,m)$ in (18) are parameterized as follows

$$\begin{aligned}
\beta &= \min(N,T) \\
\gamma &= m-T \\
\delta &= |N-T| \\
\mu(T,m,N) &= (\frac{u_\beta}{\tau_\beta}+\frac{u_{\beta-1}}{\tau_{\beta-1}})(\frac{1}{\tau_\beta}+\frac{1}{\tau_{\beta-1}})^{-1} \\
\sigma(T,m,N) &= 2(\frac{1}{\tau_\beta}+\frac{1}{\tau_{\beta-1}})^{-1} \\
\sin^2(\gamma_\beta/2) &= (\beta+1/2) \\
&\times (2\beta+\gamma+\delta+1)^{-1} \\
\sin^2(\phi_\beta/2) &= (\beta+\delta+1/2) \\
&\times (2\beta+\gamma+\delta+1)^{-1} \\
\tau_\beta^3 &= 16(2\beta+\gamma+\delta+1)^{-2} \\
&\times \sin^{-2}(\phi_\beta+\gamma_\beta) \\
&\times \sin^{-1}(\phi_\beta)\sin^{-1}(\gamma_\beta) \\
u_\beta &= 2\log[\tan(\frac{\phi_\beta+\gamma_\beta}{2})]
\end{aligned}$$

Note that the covariance matrix $\mathbf{R_b}$ has no effect on the distribution of the eigenvalue $\lambda_1^{(D)}$. A proof of Property 6.4 can be found in [21] for $\mathbf{R_b}=\mathbf{I}_T$. We now rewrite (17) as

$$\mathbf{A}\mathbf{B}^{-1}v = \frac{\theta}{1-\theta}v \tag{19}$$

and derive a new property for the following generalized eigenproblem

$$\mathbf{A}v = \theta\mathbf{B}v \tag{20}$$

This is referred to as a single Wishart setting. From (19) and (18), we get the following property.

Property 6.5: The largest eigenvalue satisfying (20), denoted $\lambda_1^{(S)}$, is distributed according to a TW distribution when $m,N\to\infty$ as $T\to\infty$ with $m,N>T$, i.e.,

$$\mathrm{P}\left\{\frac{\log[\lambda_1^{(S)}]-\mu(T,N,m)}{\sigma(T,N,m)}\le x\right\}\to TW_{\mathbb{C}}(x)$$

REFERENCES

[1] B. Escrig, "Estimation of Collision Multiplicities in IEEE 802.11-based WLANs," in *Proc. IARIA Eleventh International Conference on Networks (ICN)*, 2012.

[2] P. V. Lang Tong, V. Naware, "Signal processing in random access," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 29–39, 2004.

[3] M. Z. A. A. F. Borgonovo, L. Fratta, "Capture division packet access: a new cellular access architecture for future PCNs," *IEEE Communications Magazine*, vol. 34, no. 9, pp. 154–162, 1996.

[4] D. L. S. Y. Park, D. Park, "On Scheduling for Multiple-Antenna Wireless Networks Using Contention-Based Feedback," *IEEE Transactions on Communications*, vol. 55, no. 6, pp. 1174–1190, 2007.

[5] P. T. M. Durvy, O. Dousse, "Self-Organization Properties of CSMA/CA Systems and Their Consequences on Fairness," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 931–943, 2009.

[6] S. Verdú, "Multiuser Detection." Cambridge Press, 1998.

[7] R. Zhang, N. D. Sidiropoulos, and M. K. Tsatsanis, "Collision Resolution in Packet Radio Networks Using Rotational Invariance Techniques," *IEEE Transactions on Communications*, vol. 50, no. 1, pp. 146–155, 2002.

[8] B. Özgül and H. Deliç, "Wireless Access with Blind Collision-Multiplicity Detection and Retransmission Diversity for Quasi-Static Channels," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 858–867, 2006.

[9] A. P. R. Lin, "A new wireless network medium access protocol based on cooperation," *IEEE Transactions on Signal Processing*, vol. 53, no. 12, 2005.

[10] R. Z. G. Dimic, N.D. Sidiropoulos, "Medium access control - physical cross-layer design," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 40–50, 2004.

[11] P. X. Zheng, Y. J. Zhang, and S. C. Liew, "Multipacket Reception in Wireless Local Area Networks," in *Proc. IEEE International Conference on Communications (ICC)*, 2006.

[12] W. L. Huang, K. B. Letaief, and Y. J. Zhang, "Cross-Layer Multi-Packet Reception Based Medium Access Control and Resource Allocation for Space-Time Coded MIMO/OFDM," *IEEE Transactions on Wireless Communications*, vol. 7, no. 9, pp. 3372–3384, 2008.

[13] S. Talwar, M. Viberg, and A. Paulraj, "Blind Separation of Synchronous Co-Channel Digital Signals Using an Antenna Array. Part I. Algorithms," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1184–1197, 1996.

[14] P. Stoica and Y. Selèn, "Model-Order Selection : A review of information criterion rules," *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36–47, 2004.

[15] Q. Ding and S. Kay, "Inconsistency of the MDL: On the Performance of Model Order Selection Criteria with Increasing Signal-to-Noise Ratio," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 1959–1969, 2011.

[16] R. R. Nadakuditi and A. Edelman, "Sample Eigenvalue Based Detection of High-Dimensional Signals in White Noise Using Relatively Few Samples," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 2625–2638, 2008.

[17] M. Wax and T. Kailath, "Detection of Signals by Information Theoretic Criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.

[18] M. S. Bartlett, "A note on the multiplying factors for various χ^2 approximations," *J. R. Stat. Soc.*, vol. 16, p. 296298, 1954.

[19] D. N. Lawley, "Tests of significance of the latent roots of the covariance and correlation matrices," *Biometrica*, vol. 43, p. 128136, 1956.

[20] I. M. Johnstone, "High Dimensional Statistical Inference and Random Matrices," in *Proceeding of the International Congress of Mathematicians*, 2006.

[21] ——, "Multivariate analysis and Jacobi ensembles: Largest eigenvalue, TracyWidom limits and rates of convergence," vol. 36, no. 6, pp. 2638–2716, 2008.

[22] R. R. Nadakuditi and J. W. Silverstein, "Fundamental Limit of Sample Generalized Eigenvalue Based Detection of Signals in Noise Using Relatively Few Signal-Bearing and Noise-Only Samples," *IEEE Transactions on Signal Processing*, vol. 4, no. 3, pp. 468–480, 2010.

[23] P. O. Perry and P. J. Wolfe, "Minimax Rank Estimation for Subspace Tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 3, pp. 504–513, 2010.

[24] J. Baik, G. B. Arous, and S. Péché, "Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices," *The Annals of Probability*, vol. 33, no. 5, pp. 1643–1697, 2005.

[25] G. Xu, R. H. Roy, and T. Kailath, "Detection of number of sources via exploitation of centro-symmetry property," *IEEE Transactions on Signal Processing*, vol. 42, no. 1, pp. 102–112, january 1994.

[26] P. Stoica and M. Cedervall, "Detection tests for array procesing in unknown correlated noise fields," *IEEE Transactions on Signal Processing*, vol. 45, no. 9, september 1997.

[27] E. Fishler and H. V. Poor, "Estimation of the number of sources in unbalanced arrays via information theoretic criteria," *IEEE Transactions on Signal Processing*, vol. 53, no. 9, pp. 3543–3553, september 2005.

[28] "IEEE 802.11n standard, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 5: Enhancements for Higher Throughput," IEEE Computer Society, Tech. Rep., 2009.